International Journal of Advanced Research in Engineering Technology & Science

ISSN: 2349-2819

Email: editor@ijarets.org Volume-9, Issue-10 October - 2022 www.ijarets.org

DETECTION OF BRAIN TUMOR USING MACHINE LEARNING

Bhawna Sachdeva
Research Scholar DTU, Delhi

Prof. J. Panda
Professor DTU Delhi

ABSTRACT

At this time, tumours are the leading cause of cancer worldwide, accounting for more cases of the disease than any other risk factor. As a result of the illness, the lives of a sizeable portion of people who are battling cancer are currently in peril. The field of medicine requires a method that is quick, automated, efficient, and dependable in order to diagnose tumours such as brain tumours. The diagnosis of cancer in its earliest stages is one of the most crucial initial steps in the fight against the disease. If a tumour is found at an early enough level, medical personnel have the power to save a patient who is in immediate danger of dying if the tumour is found at an early enough stage. In order to do what it set out to do, this application makes use of a wide range of image processing methods from which to choose. It is possible for medical professionals to use this software to deliver appropriate treatment to patients who are afflicted with tumours. If successful, this treatment could lead to the survival of a significant number of these people. A tumour is nothing more than a collection of cells that have grown in an unregulated manner. Cell growth is what gives a tumour its characteristic appearance. It is not impossible for cancer cells in the brain to grow and spread in such a way that they eventually consume all of the resources that were supposed to go to healthy cells and tissues, which would result in the death of the brain. This would be a catastrophic outcome. The resulting effect would be as described above. At this time, medical personnel look at photos taken with magnetic resonance imaging (MRI) of a patient's brain in order to manually evaluate the location of a brain tumour in the patient as well as the volume of the tumour. Because of this, it is not feasible to make an accurate diagnosis of the malignancy, and the process is regarded to be one that takes place over an exceptionally extended length of time. A clump of tissue that grows in an unregulated manner and ultimately becomes uncontrollable is referred to as a tumour. Tumors can occur in almost any organ in the body. CNN, which stands for "Convolution Neural Network" and is synonymous with "Neural Network," as well as VGG 16 are examples of Deep Learning systems (visual geometry group) Transfer learning is a technique that could be used to diagnose brain tumours. Deep learning can be seen in both of these different types of technology. The success of the model is measured by how correctly it can determine whether or not a tumour is present in a particular image. This capability is used to evaluate the model's performance. The answer is yes if there is a tumour present; however, the answer is not yes if there is not a tumour present.

Keywords: *Brain, tumor, machine learning*

INTRODUCTION

The human body is comprised of a variety of organs, with the brain being the organ that is regarded as being the organ that plays the most vital and crucial role among these organs. One of the most common conditions that could lead to dysfunction in the brain is the presence of a tumour in that region of the head. A tumour is nothing more than a collection of cells that have grown in an unregulated manner. Cell growth is what gives a tumour its characteristic appearance. It is not impossible for cancer cells in the brain to grow and spread in such a way that

they eventually consume all of the resources that were supposed to go to healthy cells and tissues, which would result in the death of the brain. This would be a catastrophic outcome. The resulting effect would be as described above. At this time, medical personnel look at photos taken with magnetic resonance imaging (MRI) of a patient's brain in order to manually evaluate the location of a brain tumour in the patient as well as the volume of the tumour. Because of this, it is not feasible to make an accurate diagnosis of the malignancy, and the process is regarded to be one that takes place over an exceptionally extended length of time. Brain cancer is a disease that has the potential to be lethal and claims the lives of a considerable number of individuals on an annual basis. Because the technology required for the detection and classification of brain tumours is now readily available, it is now feasible to make a diagnosis of brain tumours at a more advanced stage in the disease's progression. The categorization of the many forms of cancer is one of the most difficult challenges faced in the field of clinical diagnosis. This study is concerned with such a system, which makes use of computer-based techniques to detect tumour blocks and classify the type of tumour using Convolution Neural Network Algorithm for MRI pictures of various patients, as well as to detect tumour blocks and classify the type of tumour. Additionally, this study aims to detect tumour blocks and classify the type of tumour. This research specifically focuses on developing a system that is able to detect tumour blocks and classify the type of tumour. MRI scans of cancer patients who have been treated with radiation therapy are analysed using a variety of image processing methods, such as image segmentation, image augmentation, and feature extraction, with the intention of locating any brain tumours that may have developed as a result of the treatment. An strategy that can identify a brain tumour through a sequence of four separate processes has been developed through the utilisation of a number of different image processing approaches. This method is broken down into four distinct stages, which are the pre-processing of images, the segmentation of images, the extraction and categorization of features, and the classification stage. In the process of locating and categorising brain tumours in magnetic resonance imaging (MRI) photographs, it is possible to achieve increased performance by employing image processing and neural network approaches. This makes it possible to gain better results.

OVERVIEW OF BRAIN AND BRAIN TUMOR

It is generally agreed that the human brain is the part of the nervous system that plays the most important role. This structure is afforded protection by the skull due to its placement within the human head, which acts as its natural environment. The human brain's primary job is to serve as the body's central control centre, directing and coordinating the activities of the body's various other organs. The human body has a particular kind of organ that enables humans to adjust to and thrive in a wide variety of different environmental situations. The human brain not only enables individuals to communicate their ideas and emotions, but it also makes it possible for them to carry out their daily activities. The investigation of the brain's anatomy, which is discussed in this section, is very necessary in order to acquire a knowledge of the fundamental concepts.

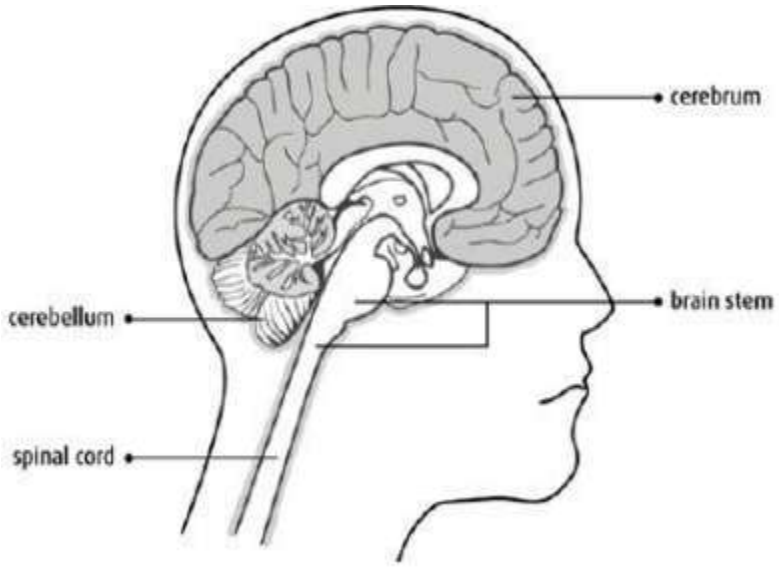

Fig.1: Basic Structure of human brain

Basic brain tumours, which are sometimes referred to as benign tumours, and secondary brain tumours are the two primary categories that are utilised for classifying brain cancers. Primary brain tumours can occur in both adults and children (malignant tumor). Gliomas are a type of brain tumour that is considered to be benign. They are composed of only one type of cell and grow very slowly in the brain. Additionally, gliomas are the most prevalent type of brain tumour that can be found. An astrocyte, which is not the same thing as a neuron, is the type of brain cell from which it originates. However, primary tumours put a large deal of pressure on the brain, which causes the brain to stop operating normally. In general, primary tumours are not as harmful as secondary tumours, but they do cause the brain to stop functioning normally. The secondary tumours are far more harmful than the primary ones since they are more malignant and can rapidly spread to other tissues. A brain tumour that originates in another part of the body is referred to as having a secondary location in the brain. Metastatic tumours are characterised by the existence of cancer cells within the body that have the potential to spread to other parts of the body, such as the brain, the lungs, or other organs. This characteristic is what gives metastatic tumours its name. The likelihood that a secondary brain tumour will become malignant is quite significant. The lung cancer, kidney cancer, bladder cancer, or bladder cancer, along with other forms of cancer, are the most prevalent causes of secondary brain tumours. Other types of cancer can also produce secondary brain tumours.

MAGNETIC RESONANCE IMAGING (MRI)

Raymond V. Damadian is credited with being the first person to make a magnetic picture in the year 1969. In 1977, magnetic resonance imaging (MRI), now widely regarded as the most cutting-edge imaging technique, was initially used to make images of the human body. Because of magnetic resonance imaging (MRI), we are able to see clear images of the complex inner workings of the brain. Furthermore, by extrapolating from those images, we are able to investigate the various types of tissues that comprise the human body. The images that are produced by magnetic resonance imaging (MRI) are of greater quality when contrasted with the findings of other medical imaging procedures, such as X-rays and CT scans. The magnetic resonance imaging (MRI) test is a tried-and-true method for determining whether or not a person has a brain tumour. T1 weighted, T2 weighted,

and FLAIR (Fluid attenuated inversion recovery) weighted MRI scans are some examples of the different types of images that can be used for tracking tumor-induced change. Another type of image that can be used is FLAIR (Fluid attenuated inversion recovery) weighted MRI scans.

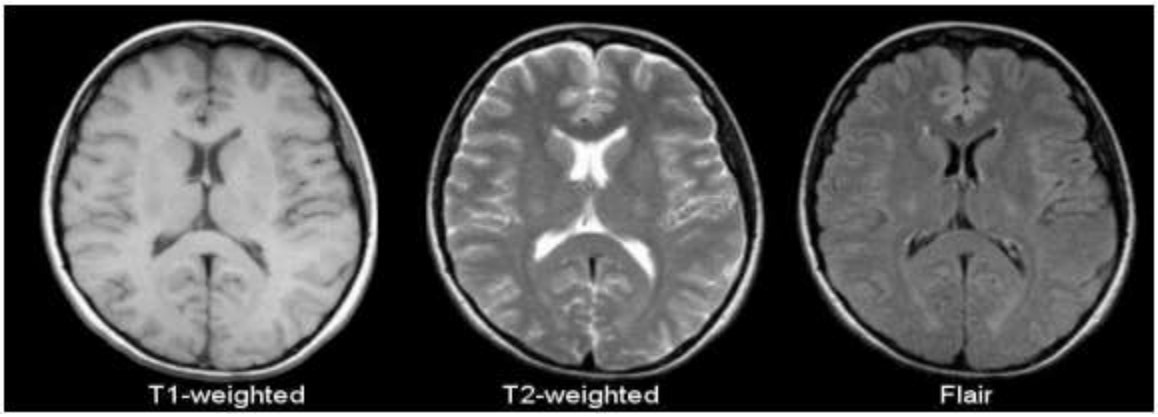

Fig 2: T1, T2 and Flair image

The bulk of magnetic resonance imaging (MRI) scans are comprised of T1-weighted and T2-weighted sequences. There is only one type of tissue that is brilliant FAT when T1 is weighted, but when T2 is weighted, there are two types of tissue that are both brilliant FAT and water. When T1 is weighted, there is only one type of tissue that is brilliant FAT. When T1 is weighted, the repetition time, also known as TR, is very short. On the other hand, when T2 is weighted, both TE and TR are quite a bit longer. The pulse sequence is comprised of a number of parameters, two of which are the repetition time (also referred to as TE) and the time to echo (also referred to as TR). Milliseconds are the units that can be used to measure each of these (ms). The amount of time that elapses between each TE repeating sequence of pulse and echo is referred to as TR. This time interval is denoted by the abbreviation. This idea is depicted in the image that can be found to the right. The amount of time that has passed after the centre of the RF pulse to the centre of the echo is what is referred to as the echo time.

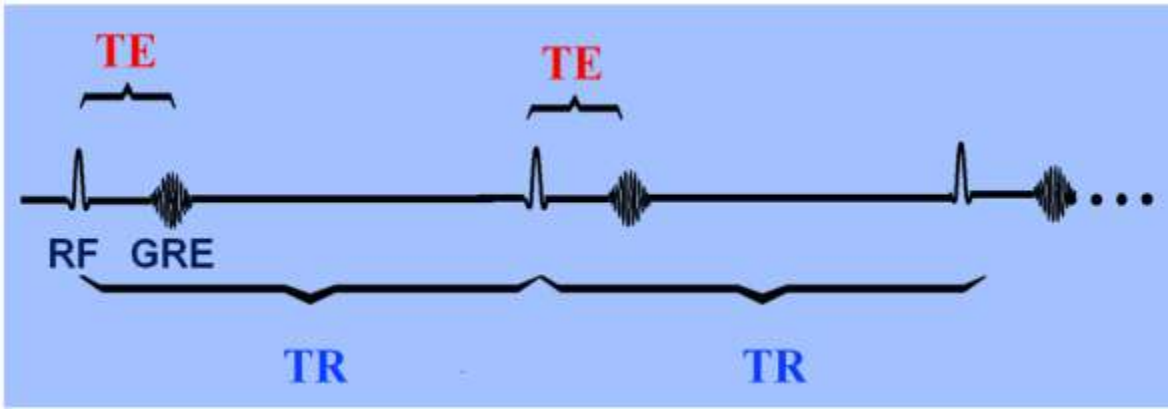

Fig. 3: Graph of TE and TR

The third sequence that is utilised rather frequently in the FLAIR. It can be seen that the Flair sequence and the T2-weighted picture are virtually indistinguishable from one another. The only difference is that the total TE

time is substantially longer than the whole TR duration. The table provides an estimate of how long their TR and TE times will be respectively.

	TR (msec)	TE (msec)
T1-Weighted (short TR and TE)	500	14
T2-Weighted (long TR and TE)	4000	90
Flair (very long TR and TE	9000	114

Fig.4: Table of TR and TE time

OBJECTIVE

[1] To provide doctors good software to identify tumor and their causes.
[2] Study on brain tumor using machine learning

Materials and methods

The performance of the two methods, ANN and CNN, in categorising the image is compared after they have both been applied to the dataset including images of brain tumours. The following procedures are carried out in order to deploy ANN on the brain tumour dataset:

1. First, the necessary software packages need to be imported into your operating system.
2. To import the data, select the folder that contains the data on your computer and click the Import button.
3. After the photographs have been read into the Data Frame, the labels for each image have been supplied (Set Image with Brain Tumor as 1, and Image without Brain Tumor as 0), and the images have then been placed in the Data Frame.
4. Read each photo one at a time so that the combined dimensions of all the photographs are brought down to 256 pixels by 256 pixels.
5. Restore the image to the original dimensions it had when it was initially saved.
6. The data set needs to be partitioned into three different parts: the train set, the validation set, and the test set.
7. Assemble a working model of the product.
8. Put the model together to create its finished appearance.
9. Place the model on the railway set in such a way that it can be operated by using the set.
10. In order to assess it, you will need to employ the model on the information contained in the test set.

The ANN model used for the purpose of this inquiry is comprised of seven different layers in total. The flatten layer is the first layer in the process, and its purpose is to convert the 256x256x3 images into a single-dimensional array. This layer is located at the beginning of the process. Following the initial level,

the next five levels are dense layers, with the activation function relu and the number of neurons in each layer sequentially being 128, 256, 512, 256, and 128. The next tier is a substantial one and comes after these five levels. A relu value is used to indicate the activation function of each individual layer. These five layers are what are known as the hidden layers, and the final dense layer, which is the one that acts as the output layer and has an activation function that is sigmoid, is what serves as the hidden layer. Within the layers that are hidden from view, there is one neuron that is representative of each of the two categories. In order to achieve the desired result, it is built with the Adam optimization technique and a binary crossentropy loss function. These are the two components that make up the loss function. The validation images and training photographs are used to assist in the generation of the model, which is then trained using those photographs. After the training session, the model is evaluated with the use of the test photo collection in order to see how well it performed. After that, the processing of the same dataset is handed over to an algorithm known as CNN. When applying CNN to the dataset that contained information on brain tumours, the following steps were carried out:

1. Into your operating system, import the required software packages.
2. Select the data folder that is located on your computer and import it (Yes and No)
3. Label the photographs using the appropriate image classes (1 for Brain Tumor and 0 for No Brain Tumor)
4. Transform the images into a form that can be recognised (256X256)
5. Make the image's sizes more consistent with one another.
6. Separate the photographs into three categories: the images of the train, the images of the validation, and the images of the test set.
7. Construct a chronological representation of the data.
8. Ensure that the model's compilation is finished.
9. Test it out on the train dataset and see what results it produces (use validation set to evaluate the training performance).
10. Determine the overall performance of the model by analysing the test images.
11. Produce a graph that contrasts the precision of the training data with that of the validation data.
12. Step 12 entails drawing the confusion matrix to compare the actual output with the planned result.

After the application of numerous layers, the initial picture is reduced in size to 256 by 256 pixels, which ultimately results in the construction of the CNN sequential model. The input picture has the convolve layer applied to it with the relu set as the activation function, and the padding is set to the same value as the input image. This results in output images that are similar to the input image. There are 32 filters, 32 filters, 64 filters, 128 filters, and 256 filters in each of the convolve layers, accordingly. In this particular illustration, the max pooling method is applied in conjunction with the 2x2 window size, and the droupout function is invoked with 20 percent of the droupout as its argument. The flatten method is used in order to reduce the number of dimensions present in the features array to a single value. This is accomplished by using the flatten strategy. The dense method, which contains a total of 256 units, is used to construct the entirely connected layer. The activation function that is used is referred to as relu, and it is used in conjunction with the dense method. After that, the completely linked layer is applied. In the output layer, there is a sigmoid function that acts as an activation function and one unit that corresponds to each of the two classes that are being predicted. Additionally, there is a weighted average of the output layer's output units. Figure 5 provides a visual

representation of the fundamental architecture of the CNN model. Python is utilised as a programming language in order to carry out the implementation, which is carried out in Google Colab. A total of 200 epochs spanning a century and a half are used in the execution of the model, during which both the training datasets and the validation datasets are used. The history of the execution is stored and exhibited so that the models that are produced may be appreciated in greater depth. This allows for improved accuracy.

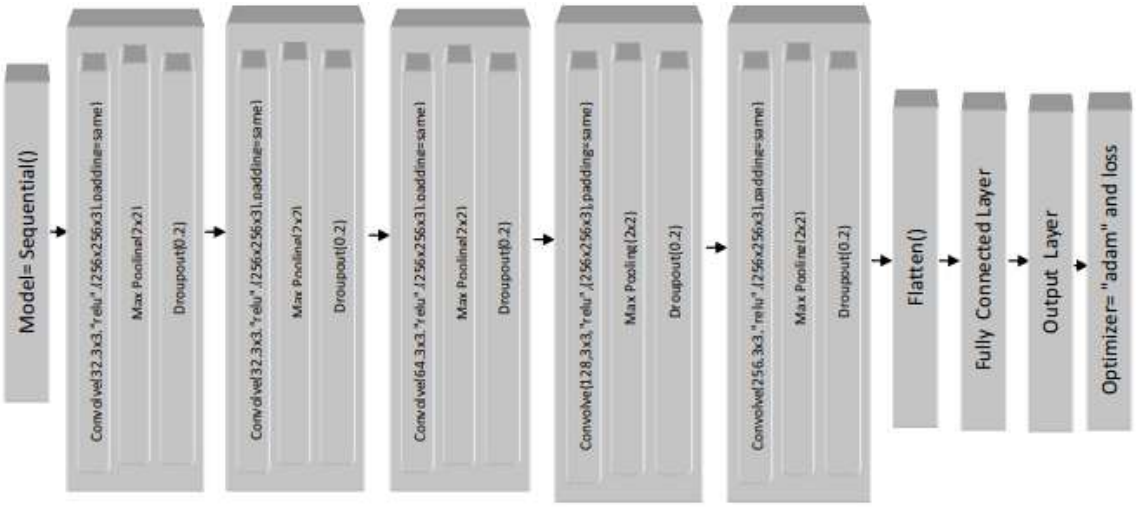

Figure 5. Architecture of CNN model

Experimental Result Analysis

A variable known as data, which stores information of the ndarray data type, is given the data that pertains to the images that are to be displayed (ndarray data structure). The labels for the photographs' corresponding class categories are also generated and saved in the variable data target, which is also an ndarray. These labels are stored in the variable data target. The photographs are then loaded into the data frame after that has been completed. The image dataset is divided into three distinct parts: the training dataset, the validation dataset, and the testing dataset. Each of these parts is referred to as a dataset. The accuracy and loss that were obtained by applying the ANN model to the training and validation datasets are depicted in Figure 6. These results can be compared to one another in a side-by-side format. The results of applying the ANN model to the training data for fifty epochs on the training data set produce an accuracy of 97.13 percent for training and 71.51 percent for validation respectively. When the same algorithm is utilised on the data obtained from the testing, it produces an accuracy of 80.77 percent. The accuracy of the training is significantly higher than the accuracy of the validation.

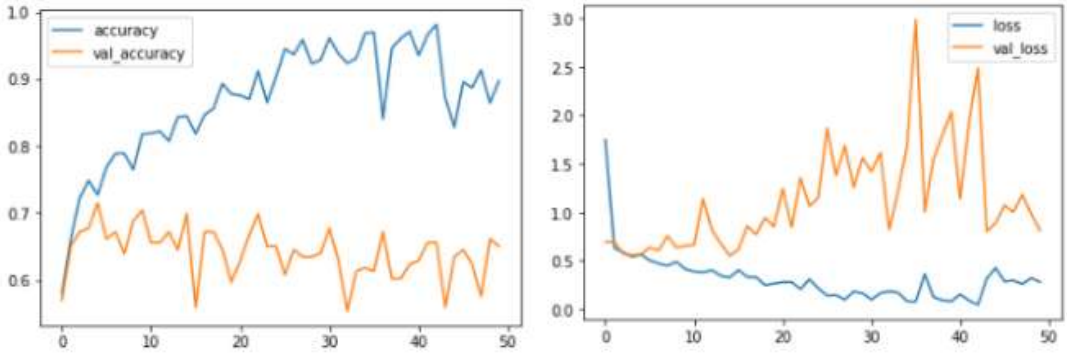

Figure 6. Comparing training/validation accuracy and loss of ANN model

The model has reached the highest possible accuracy for validation, which is 94.000 percent, after being trained and validated on the training dataset for a total of 200 iterations. The training dataset was used to develop the model in the first place. The relationship between the training accuracy and the validation accuracy is depicted in the plot that follows in figure 7, which can be seen below. Additionally, the relationship between the training loss and the validation loss is also depicted in this plot. Both of these relationships can be seen below.

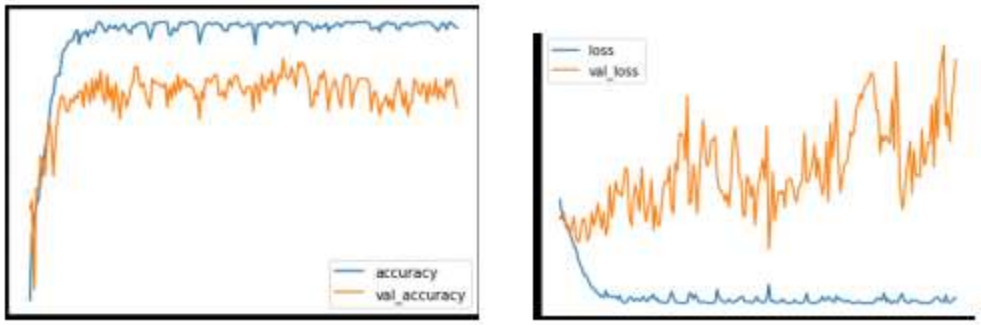

Figure 7. Comparing training/validation accuracy and loss of CNN model

Applying the test image dataset allows for accurate assessment of the model The confusion matrix for the output that was anticipated can be found in Figure 8, which can be found below. The results of creating a forecast for what will be tested and validated are shown down below.

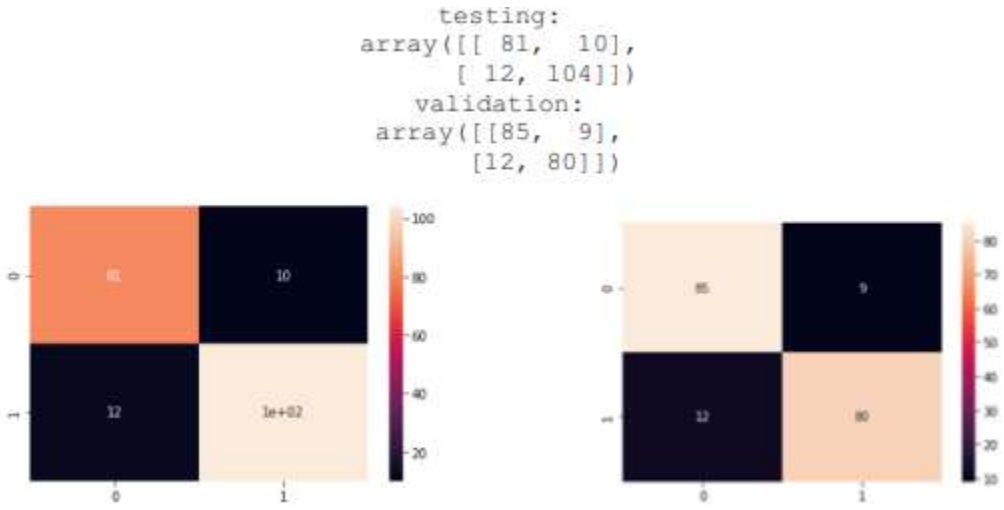

Figure 8. Confusion matrix for the testing and validation dataset using CNN model

The precision, recall and the f1-score of both the models are given in Figure 9.

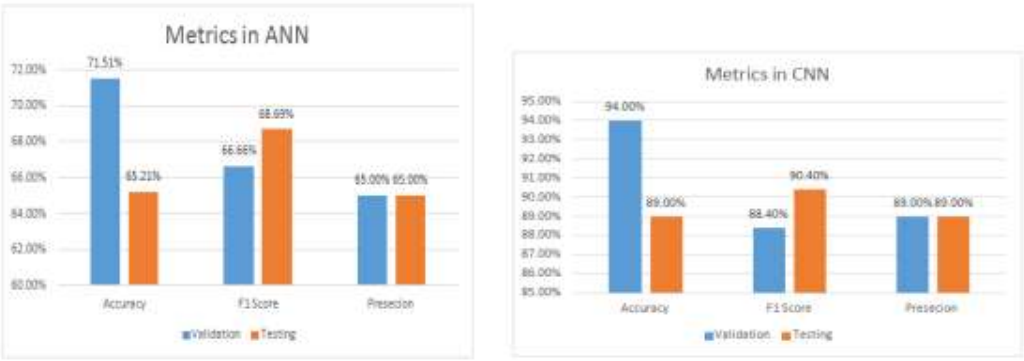

Figure 9. Metrics

The CNN model achieves an accuracy of 89 percent in terms of its application of the data when making use of the data from the tests. If we look at the scores for accuracy, recall, and f1, as well as compare them to the performance of ANN and CNN in determining whether or not a brain tumour exists, it becomes abundantly clear that CNN is the superior supporting strategy due to the fact that it has the highest precision value. This is the case because CNN is able to determine whether or not a brain tumour exists.

Conclusion

When looking at image datasets, it is common knowledge that CNN is one of the methods that has the potential to produce the most accurate results possible. Because CNN is able to reduce the size of the image without sacrificing the information that is necessary in order to generate predictions, the network is in a position to provide forecasts. The ANN model that was built for this research project achieves a testing accuracy of 65.21 percent; however, this accuracy could be improved by incorporating more picture data into the algorithm. It is possible to achieve the same result by utilising a variety of photo enhancement techniques; following this step, one is able to evaluate how well ANN and CNN models performed. The model that is described in this article

was developed through a process that included both successful and unsuccessful attempts at the same thing. It is possible that in the not-too-distant future, optimization strategies will be used to determine the possible number of layers and filters that will be incorporated into a model. It would be an interesting turn of events if something like this actually took place. According to the information that has been provided, the CNN method gives the impression of being the best method that can be used in order to make a diagnosis about the possible presence of a brain tumour. This conclusion is based on the information that has been presented.

References

[1] Raheleh Hashemzehi Seyyed Javad Seyyed Mahdavi Maryam Kheirabadi Seyed Reza Kamel 2020 Detection of brain tumors from MRI images base on deep learning using hybrid model CNN and NADE Elsevier B.V. on behalf of Nalecz Institute of Biocybernetics and Biomedical Engineering of the Polish Academy of Sciences Online Publication.

[2] Javeria Amin Muhammad Sharif Mudassar Raza Mussarat Yasmin 2018 Detection of Brain Tumor based on Features Fusion and Machine Learning Journal of Ambient Intelligence and Humanized Computing Online Publication.

[3] Rajeshwar Nalbalwar Umakant Majhi Raj Patil Prof.Sudhanshu Gonge 2014 Detection of Brain Tumor by using ANN International Journal of Research in Advent Technology 2.

[4] Fatih Özyurt Eser Sert Engin Avci Esin Dogantekin 2019 Brain tumor detection based on Convolutional Neural Network with neutrosophic expert maximum fuzzy sure entropy Elsevier Ltd 147.

[5] Fausto Milletari Seyed-Ahmad Ahmadi Christine Kroll Annika Plate Verena Rozanski Juliana Maiostre Johannes Levin Olaf Dietrich Birgit Ertl-Wagner Kai Bötzel, Nassir Navab 2016 Hough-CNN: Deep learning for segmentation of deep brain regions in MRI and ultrasound Elsevier Inc 164 92-102.

[6] Dena Nadir George Hashem B Jehlol Anwer Subhi Abdulhussein Oleiwi 2015 Brain Tumor Detection Using Shape features and Machine Learning Algorithms International Journal of Scientific & Engineering Research 6 12 454-459.

[7] Sobhangi Sarkar Avinash Kumar Sabyasachi Chakraborty Satyabrata Aich Jong-Seong Sim Hee-Cheol Kim 2020 A CNN based Approach for the Detection of Brain Tumor Using MRI Scans Test Engineering and Management 83 16580 – 16586.

[8] Komal Sharma Akwinder Kaur Shruti Gujral 2014 Brain Tumor Detection based on Machine Learning Algorithms International Journal of Computer Applications 103 7-11.